Journal of Computational and Applied Mathematics 29 (1990) 91–100
North-Holland

A fifth-order interpolant for the Dormand and Prince Runge–Kutta method

M. CALVO, J.I. MONTIJANO and L. RANDEZ
Departamento de Matemática Aplicada, Universidad de Zaragoza, 50009 Zaragoza, Spain

Received 5 October 1988
Revised 24 January 1989

Abstract: A family of fifth-order interpolants for the fifth-order solution provided by the Dormand and Prince Runge–Kutta pair RK5(4)7M which requires two additional function evaluations per step is presented. An optimal interpolant in this family has been determined by choosing the parameters to minimize the leading coefficients of the local truncation error of the continuous solution. Some numerical experiments with the nonstiff DETEST problems show that the proposed optimal method has a good interpolatory behavior.

Keywords: Ordinary differential equations, Runge–Kutta methods, interpolation.

1. Introduction

In the last decade a number of efficient Runge–Kutta codes have been written for the numerical solution of ODEs. Two representative examples are RKF45 [10] and its successor DERKF in DEPAC [11] produced by Shampine and Watts and DVERK [7] by Hull, Enright and Jackson. The first implements a pair of formulas of orders 4 and 5 of Fehlberg [3], while the second is based on a pair of orders 5 and 6 of Verner; both do local extrapolation. Although for several years these codes have proved to be reliable and efficient, it seems likely that, in the future, a new generation of RK codes will appear.

First of all, new pairs of RK formulas which may be more efficient than those in use have been proposed. In particular Dormand and Prince [1] derived a pair RK5(4)7M of orders 4 and 5 which seems to be superior to the classical Fehlberg pair. Furthermore, in some applications dense output is required and in such a case RK methods must frequently shorten the stepsize and are therefore inefficient. Thus the new RK solvers should have the possibility of producing, if necessary, reliable approximations to the solution at any point of the integration interval without stepsize adjustment and with little additional computational cost. This has been the main reason for developing the so-called continuous or interpolatory RK methods.

Horn [5] showed how to construct RK extensions of the fourth- and fifth-order formulas due to Fehlberg with some additional function evaluations. In particular, with one extra stage she produces a fourth-order approximation $y_h(t_n + \theta h_n)$ to the solution at the point $t_n + \theta h_n$ for all $\theta \in [0, 1]$. This approximation has the form

$$y_h(t_n + \theta h_n) = y_n + h_n \sum_{j=1}^{s+1} b_j(\theta) f_j,$$

where $f_i = f(t_n + c_i h_n,\ y_n + h_n \sum_{j=1}^{i-1} a_{ij} f_j)$, $i = 1, \ldots, s+1$. Here the first s stages are given by the Fehlberg method. For any given $\theta \in (0, 1)$ she shows how to produce a fifth-order solution with two extra stages. Moreover, she proves that with five extra stages it is possible to derive fifth-order approximations for all $\theta \in [0, 1]$. A serious disadvantage of Horn's formulas is that they are not continuous, i.e., the continuous solution $y_h(t_n + \theta h_n)$ does not tend to the value y_{n+1} computed by the discrete formula as $\theta \to 1$.

Shampine [8,9] and Shampine et al. [4], using an interpolatory procedure, have proposed some continuous extensions of Fehlberg and Dormand–Prince formulas that are $\mathscr{C}^1$-globally and require only two additional stages. Taking into account that we know the solution y_n and its derivative y_n' at t_n and the fifth-order approximations at $t_{n+1} = t_n + h_n$, y_{n+1} and y_{n+1}' (this value will be necessary for the next step), Shampine shows that with one additional stage it is possible to get a fifth-order solution $y_{n+1/2}$ at the midpoint of the step, $t_n + \frac{1}{2}h_n$, and by computing $y_{n+1/2}' = f(t_n + \frac{1}{2}h,\ y_{n+1/2})$, we may do quintic Hermite interpolation to get an approximate solution accurate to $O(h_n^6)$ at any point of the interval $[t_n,\ t_{n+1}]$.

Enright et al. [2] have proposed a general "boot-strapping" procedure for the construction of interpolants of RK methods and applied this technique to develop families of interpolants for RKF45 and DVERK. They choose the free parameters to minimize the max-norm of the coefficients of the principal error term for the two formulas of each pair in the intervals $[t_n,\ t_n + h_n]$ and $[t_n,\ t_n + 2h_n]$.

The aim of this paper is to construct a fifth-order family of interpolants for the Dormand and Prince RK pair RK5(4)7M (which for simplicity will be referred to as DOPRI5(4)) using the interpolatory approach of Shampine [8,9]. Since we have approximations to the solution and its derivative at two consecutive grid points, t_n and t_{n+1}, it is sufficient to compute a fifth-order approximation to the solution and its derivative at some intermediate point $t_n + \sigma h \in (t_n,\ t_{n+1})$. It is then possible to calculate a quintic Hermite interpolating polynomial which provides a fifth-order solution for all $t \in [t_n,\ t_{n+1}]$. Obviously such a solution will be $\mathscr{C}^1$-continuous globally.

The paper is organized as follows: in the next section we show that by adding a stage to DOPRI5(4) we can get a fifth-order approximation at any point in the interval $[t_n,\ t_{n+1}]$. Since in this process there are four degrees of freedom we obtain a family of RK interpolants for DOPRI5(4) depending on four parameters. Next an optimal method is selected by choosing the free parameters to minimize in a certain norm the leading coefficients of the local truncation error of the continuous solution. In particular, with this error measure our interpolant is more accurate than an earlier interpolant (DPS) proposed by Shampine [9]. Finally in Section 3 we present some numerical results with some typical equations from the nonstiff DETEST problems to show the accuracy of the continuous solution at all points of the integration interval.

2. Construction of interpolants for the Dormand–Prince pair

Let

$$\begin{aligned} &y'(t) = f(t,\ y(t)), \quad t \geqslant t_0,\ y \in \mathbb{R}^N, \\ &y(t_0) = y_0, \end{aligned} \tag{1}$$

be the initial-value problem to be solved by the explicit 7-stage RK pair DOPRI5(4). Such a method is defined by its Butcher table of coefficients

$$\begin{array}{c|c} c & A \\ \hline & b^{*\mathrm{T}} \\ & b^{\mathrm{T}} \end{array}, \quad A \in \mathbb{R}^{7\times 7},\ b^*,\ b,\ c \in \mathbb{R}^7, \tag{2}$$

where $c = Ae$, $e = (1,\ldots,1)^{\mathrm{T}} \in \mathbb{R}^7$ and b^*, b are the coefficients of the fourth- and fifth-order solution respectively. Hence, given an approximation y_n of the solution of (1) at point t_n, two approximations y^*_{n+1} and y_{n+1} to the solution at $t_{n+1} = t_n + h$ of orders 4 and 5 respectively are given by the formulas

$$y^*_{n+1} = y_n + h\sum_{j=1}^{7} b^*_j f_j, \tag{3.a}$$

$$y_{n+1} = y_n + h\sum_{j=1}^{6} b_j f_j, \tag{3.b}$$

where

$$f_j = f\left(t_n + c_j h,\ y_n + h\sum_{i=1}^{j-1} a_{ji} f_i\right), \quad j = 1,\ldots,7. \tag{3.c}$$

Denoting by $y(t;\ t_n,\ y_n)$ the local solution of (1) at the point $(t_n,\ y_n)$, i.e., the solution of the differential equation that satisfies $y(t_n) = y_n$ and assuming that $f(t,\ y)$ is sufficiently smooth, it is well known that the local error of a RK solution y_{n+1} of order p at the point $t_n + h$ can be written in the form

$$y(t_n + h,\ t_n,\ y_n) - y_{n+1} = h^{p+1} \sum_{\rho(\tau)=p+1} C(\tau)F(\tau)(y_n) + \mathrm{O}(h^{p+2}).$$

Here τ is a rooted tree, $F(\tau)(y_n)$ the elementary differential associated with τ at the point $(t_n,\ y_n)$ and $C(\tau)$ the corresponding weight. The above sum is extended to all rooted trees of order $\rho(\tau) = p + 1$. It is worth recalling that $F(\tau)(y_n)$ depends only on the differential equation and the starting point $(t_n,\ y_n)$ while $C(\tau)$ depends only on the coefficients (2) of the formula.

We assume that the integration is advanced with the fifth-order approximation y_{n+1}, i.e., local extrapolation is done, and the fourth-order solution is employed to estimate the local error and select the stepsize according to the error tolerance provided by the user. Notice in (2) that the last row of A coincides with b^{T}, so that if the step succeeds, the last stage of one step is the same as the first stage of the next step. This means that if the stepsize selection is effective enough to make rejected steps unusual, then it is fair to say that DOPRI5(4) costs 6 evaluations per step.

As it was remarked before, our aim is to construct a fifth-order $\mathscr{C}^1$-continuous extension of the solution provided by DOPRI5(4) with minimum computational cost and with a local error as small as possible in the whole interval $[t_n,\ t_n + h]$. Clearly if we denote by $\hat{y}(t)$ such a continuous extension, it will satisfy

$$\begin{aligned} &\hat{y}(t_n) = y_n, && \hat{y}(t_{n+1}) = y_{n+1}, \\ &\hat{y}'(t_n) = f(t_n,\ y_n), && \hat{y}'(t_{n+1}) = f(t_{n+1},\ y_{n+1}). \end{aligned} \tag{4}$$

To construct the fifth-order approximation $\hat{y}(t)$ on the whole interval $[t_n,\ t_n + h]$ we follow the

interpolatory approach due to Shampine [9]. Since we have fifth-order approximations (4) at both ends of the interval $[t_n, t_n + h]$, it will be sufficient to have fifth-order approximations $y_{n+\sigma}$, $y'_{n+\sigma} = f(t_{n+\sigma h}, y_{n+\sigma h})$ at some intermediate point $t_{n+\sigma h} = t_n + \sigma h$, $\sigma \in (0, 1)$. Then the quintic Hermite polynomial $\hat{y}(t_n + \theta h)$, $\theta \in (0, 1)$, that satisfies (4) and

$$\hat{y}(t_n + \sigma h) = y_{n+\sigma h}, \qquad \hat{y}'(t_n + \sigma h) = y'_{n+\sigma h} = f(t_{n+\sigma h}, y_{n+\sigma h}), \tag{5}$$

gives a fifth-order solution for all $\theta \in (0, 1)$ and can be written in the form

$$\hat{y}(t_n + \theta h) = y_n + (\theta h) \sum \hat{b}_j(\theta) f_j, \tag{6}$$

where $\hat{b}_j$ are fifth-degree polynomials in θ and the sum is extended not only to the function evaluations of DOPRI5(4) but also to those involved in the computation of (5).

Before considering the calculation of $y_n + \sigma h$, let us recall some relevant facts of the derivation of DOPRI(4) that are important for our work. Dormand and Prince assume that the coefficients A, c satisfy

$$Ac = \tfrac{1}{2}\left(c^2 - c_2^2 e_2\right), \qquad Ac^2 = \tfrac{1}{3}\left(c^3 - c_2^3 e_2\right), \tag{7}$$

where $c^q = (c_1^q, \ldots, c_7^q)^{\mathrm{T}}$ and e_j is the j-unit canonical vector of components $(e_j)_i = \delta_{ij}$. These conditions are called simplifying assumptions and it may be verified that on assuming (7), a RK method with coefficients $b = (b_j)$ has order 5 if and only if $b_2 = 0$ and the remaining coefficients satisfy

$$\begin{aligned} &b^{\mathrm{T}} c^j = 1/(j+1), \quad j = 0, \ldots, 4, \\ &b^{\mathrm{T}}(c \cdot Ae_2) = 0, \qquad b^{\mathrm{T}} A c^3 = \tfrac{1}{20}, \qquad b^{\mathrm{T}} A e_2 = 0, \qquad b^{\mathrm{T}} A^2 e_2 = 0. \end{aligned} \tag{8}$$

Here $u \cdot v$ denotes the componentwise product of vectors u and v, i.e., $(u \cdot v)_i = u_i v_i$. In order to simplify the calculation of the fifth-order solution of DOPRI5(4), Dormand and Prince assume also that

$$b^{\mathrm{T}} A = b^{\mathrm{T}} - (b \cdot c)^{\mathrm{T}}.$$

This condition together with (7) implies that the last three equations of (8) can be eliminated and $c_6 = 1$. In this way they derive a family of fifth-order methods depending on parameters c_3, c_4, c_5 which are then chosen to minimize the principal truncation error term.

Clearly if there were $\sigma \in (0, 1)$ and $\bar{b}_j \in \mathbb{R}^7$ such that $\bar{y}_{n+\sigma} = y_n + (\sigma h)\sum_{j=1}^{7} \bar{b}_j f_j$ is an $O(h^6)$ approximation to $y(t_n + \sigma h; t_n, y_n)$, this fifth-order solution at essentially no extra cost would be a very convenient approximation to be used in (5). However, it is straightforward to verify that this is not possible, so it will be necessary to add at least one stage to obtain a fifth-order solution. Thus we consider an 8-stage RK method with the table of coefficients given by

$$\begin{array}{c|c} c & A \\ c_8 & a^{\mathrm{T}} \\ \hline & b'^{\mathrm{T}} \end{array} = \begin{array}{c|c} c' & A' \\ \hline & b'^{\mathrm{T}} \end{array}, \tag{9}$$

obtained by adding a final stage to DOPRI5(4). As usual we assume $c_8 = a^{\mathrm{T}} e$ where $e = (1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^8$. We want to find the relations that the new parameters a and b' must satisfy so that

$$y_{n+\sigma} = y_n + (\sigma h) \sum_{j=1}^{8} b'_j f_j, \tag{10}$$

is a fifth-order approximation to $y(t_n + \sigma h,\ t_n,\ y_n)$. Because the simplifying assumptions allow a considerable reduction of the number of fifth-order conditions, let us assume that the new matrix A' and the vector c' also satisfy these assumptions. Since the submatrices A and c of A' and c' already satisfy (7) it is straightforward to verify that the simplifying assumptions hold for A' and c' if and only if

$$a^{\mathrm{T}}c' = \tfrac{1}{2}c_8^2 \qquad a^{\mathrm{T}}c'^2 = \tfrac{1}{3}c_8^3. \tag{11}$$

Now doing the same calculations that permitted us to reduce the fifth-order conditions of (3.b) to the form (8), it can be proved that (10) is a fifth-order formula if and only if $b_2' = 0$ and

$$b'^{\mathrm{T}}c'^{j} = \sigma^j/(j+1), \quad j = 0, \ldots, 4,$$
$$b'^{\mathrm{T}}(c' \cdot A'e_2) = 0, \qquad b'^{\mathrm{T}}A'c'^3 = \tfrac{1}{20}\sigma^4, \qquad b'^{\mathrm{T}}A'e_2 = 0, \qquad b'^{\mathrm{T}}A'^2e_2 = 0. \tag{12}$$

To study the compatibility of (11) and (12) we introduce the new scalar variables $\gamma = b_9'a^{\mathrm{T}}c'^3$, $\gamma_1 = -b_8'a^{\mathrm{T}}Ae_2$, $\gamma_2 = -b_8'a^{\mathrm{T}}e_2$. Then taking into account that $c_6 = c_7 = 1$ and denoting by $(x)_{i-j}$ the row vector $(x_i, \ldots, x_j)$, equations (11) and (12) can be rewritten equivalently in the form

$$\begin{pmatrix} c_{3-7} & c_8 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ c_{3-7}^4 & c_8^4 & 0 & 0 \\ (Ac^3)_{3-7} & 0 & 0 & 0 \\ (A^2e_2)_{3-7} & 0 & -1 & 0 \\ (Ae_2 \cdot c)_{3-7} & 0 & 0 & -c_8 \\ (Ae_2)_{3-7} & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} b_3' \\ \vdots \\ b_6' \\ b_7' \\ b_8' \\ \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sigma \\ \vdots \\ \frac{1}{5}\sigma^4 \\ \frac{1}{20}\sigma^4 - \gamma \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{13}$$

$$\begin{pmatrix} c_2 & c_3 & c_4 & c_5 & 1 \\ c_2^2 & c_3^2 & c_4^2 & c_5^2 & 1 \\ c_2^3 & c_3^3 & c_4^3 & c_5^3 & 1 \\ 0 & a_{32} & a_{42} & a_{52} & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{82} \\ a_{83} \\ a_{84} \\ a_{85} \\ a_{87} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}c_8^2 - a_{86} \\ \frac{1}{3}c_8^3 - a_{86} \\ \gamma/b_8' - a_{86} \\ -\gamma_1/b_8' - a_{86}a_{62} \\ -\gamma_2/b_8' \end{pmatrix}, \tag{14}$$

$$b_1' = 1 - \sum_{j=3}^{8} b_j', \qquad b_2' = 0, \qquad a_{81} = c_8 - \sum_{j=2}^{7} a_{8j}. \tag{15}$$

An elementary but tedious calculation shows that the matrix of coefficients in (13), which depends only on the parameter c_8, is singular only for $c_8 = 0$, $c_8 = 1$, and $c_8 = c_8^*$, where $c_8^* = 0.91661200243235649\ldots$ is the real root of the polynomial $P(x) = -49950x^3 + 70945x^2 - 26322x + 2988$. Furthermore, for the c_i-values of DOPRI5(4) ($c_2 = \frac{1}{5}$, $c_3 = \frac{3}{10}$, $c_4 = \frac{4}{5}$, $c_5 = \frac{8}{9}$) the constant matrix in (14) is nonsingular. In view of these observations we proceed as follows to determine the parameters of fifth-order solutions:

(i) Choose $\sigma \in (0, 1)$, $c_8 (\neq 0, 1, c_8^*)$, γ and a_{86}.
(ii) Compute b_3', b_4', b_5', b_6', b_7', b_8', γ_1, γ_2 from (13).
(iii) Compute a_{82}, a_{83}, a_{84}, a_{85} and a_{87} from (14).
(iv) Compute b_1' and a_{81} from (15).

Once such a fifth-order solution $y_{n+\sigma}$ has been obtained, with another function evaluation we may compute $y'_{n+\sigma} = f(t_{n+\sigma}, y_{n+\sigma})$ and by Hermite interpolation we have the fifth-order solution (6) in the whole interval $[t_n, t_n + h]$.

Note. Taking the values $c_8 = \sigma = \frac{1}{2}$, $a_{86} = 0$, $\gamma = -0.003003188$ we have the fifth-order interpolant DPS proposed by Shampine in [9].

Since we have a family of fifth-order methods with four degrees of freedom (σ, c_8, γ, a_{86} with $\sigma \in (0, 1)$ and $c_8 \neq 0, 1, c_8^*$), for each set $\mu = (\sigma, c_8, \gamma, a_{86})$ we may define by interpolation a RK continuous solution $y_\mu(t_n + \theta h)$, $\theta \in [0, 1]$ whose local error will be given by

$$y(t_n + \theta h, t_n, y_n) - y_\mu(t_n + \theta h) = (\theta h)^6 \sum_{\rho(\tau)=6} C_{\theta,\mu}(\tau) F(\tau)(y_n) + O(h^7),$$

where the weights $C_{\theta,\mu}(\tau)$ depend on θ and the free parameters. Now we define as a measure of the local error of this formula the quantity

$$g_\mu^* = \int_0^1 g_\mu(\theta)\, d\theta, \tag{16}$$

where

$$g_\mu(\theta) = \theta^6 \sqrt{\frac{\sum_{\rho(\tau)=6} \{C_{\theta,\mu}(\tau)\}^2}{\sum_{\rho(\tau)=6} \{C_{1,\mu}(\tau)\}^2}}\,. \tag{17}$$

Note that (17) represents the ratio of the l_2-norms of the weights of the sixth-order elementary differentials at the θ-point and at the end point of the interval $[t_n, t_n + h]$. Because at the end point of the interval we have for all μ the fifth-order solution of DOPRI5(4), the coefficients $C_{1,\mu}(\tau)$ are independent of μ and the denominator in (17) is the constant $3.99 \cdot 10^{-4}$ previously calculated by Dormand and Prince [1].

Next we consider how to choose the free parameters in order to get a continuous method with minimal local error in the sense of (16). This minimization process has been carried out numerically in the following way: First a grid was established in the space of parameters, excluding points close to some undesirable values of the parameters and g_μ^* was computed on the points of this grid. As a consequence of this search some grid points close to the minimum were located. Taking them as starting values, better values were found by a descent method. Finally we took simple rational values close to the computed optimal ones because they are more convenient from a computational point of view. In this way we found the "optimal" values

$$c_8 = \tfrac{2}{5}, \qquad \sigma = \tfrac{2}{5}, \qquad \gamma = -\tfrac{1}{250}, \qquad a_{86} = \tfrac{1}{20}, \tag{18}$$

for which $g_\mu^* = 0.68$. From (13), (14) and (15) the coefficients of the additional stage and the new fifth-order solution are

$$a_{81} = -\tfrac{24018683}{8152320000}, \quad a_{82} = \tfrac{25144}{43425}, \quad a_{83} = -\tfrac{76360723}{337557000},$$
$$a_{84} = \tfrac{349808429}{2445696000}, \quad a_{85} = -\tfrac{13643731773}{144024320000}, \quad a_{86} = \tfrac{1}{20}, \quad a_{87} = -\tfrac{12268567}{254760000},$$
$$b'_1 = \tfrac{2104901}{9204000}, \quad b'_2 = 0, \quad b'_3 = \tfrac{27162112}{21341775}, \quad b'_4 = \tfrac{134233}{920400},$$
$$b'_5 = -\tfrac{13268529}{162604000}, \quad b'_6 = \tfrac{13486}{402675}, \quad b'_7 = -\tfrac{3162}{95875}, \quad b'_8 = -\tfrac{1737}{3068}.$$

The value of $g_\mu^* = 0.92$ corresponding to the continuous extension given by Shampine [9] is considerably larger than the value of our method. In Fig. 1 we have plotted the graphs of the

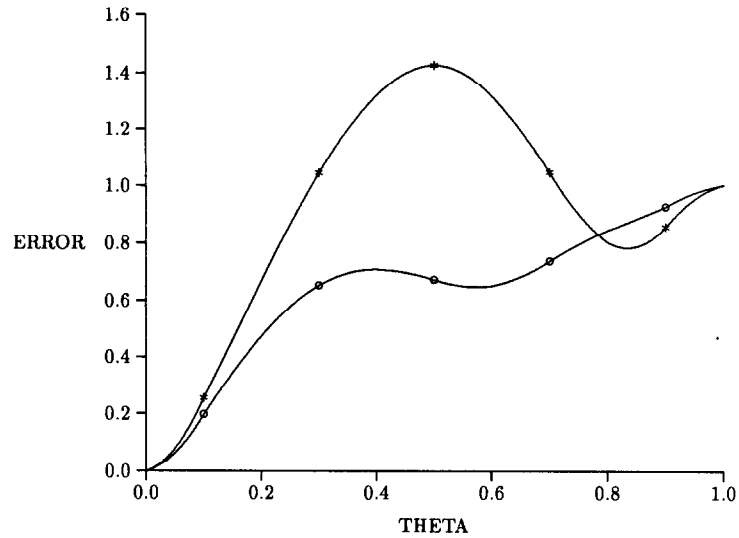

Fig. 1.

functions $g_\mu(\theta)$ (denoted by ERROR) as functions of θ for the DPS method of Shampine (in the figure, $*$ —— $*$) as well as our method (in the figure, $\circ$——$\circ$). Clearly the behavior of $g_\mu(\theta)$ for our method is better than that of Shampine's method. Although our optimal interpolant has been chosen to minimize the functional (16), other functionals could be chosen and our minimization procedure applied in the same way. In particular we have also considered the functional

$$\bar{g}_\mu(\theta) = \max\left\{ \frac{\theta^6 C_{\theta,\mu}(\tau)}{\alpha(\tau)} \middle| \rho(\tau) = 6,\ \theta \in [0, 1] \right\},$$

where $\alpha(\tau)$ is the number of times that an elementary differential appears in the Taylor expansion of the local solution. This amounts to using the max-norm instead of the l_2-norm to measure the coefficients of the sixth-order elementary differentials. It is remarkable that the values (18) are also nearly optimal for this other functional. Enright et al. [2] have constructed and analyzed families of interpolants using a different approach. The number of free parameters are generally smaller than in our procedure. In particular, for the fifth-order continuous extension of the Fehlberg pair they have only two free parameters while in our approach we have four parameters. On the other hand, they have selected an optimal method so that the four functionals $\eta_{j,p}$ for $j = 1, 2$ and $p = 4, 5$ given by

$$\eta_{j,p} = \max\left\{ \frac{\theta^6 C_{\theta,\mu}(\tau)}{\alpha(\tau) p!} \middle| \rho(\tau) = p,\ \theta \in [0, j] \right\},$$

are "small".

3. Numerical experiments

Whenever a new method for the numerical solution of nonstiff ODEs is proposed it seems to be customary to study its behavior for the classical collection of nonstiff DETEST problems and

to compare the new method with other existing methods of the same order and computational cost. For brevity we show here the results with some typical DETEST problems whose exact solution can be easily computed (A1, A2, A3, A4, D2, D3, D4, D5) and the simple nonlinear problem

$$y' = \left(y + \sqrt{t^2 + y^2}\right)/t, \quad t \in [1, 20],$$
$$y(1) = 0, \tag{19}$$

whose exact solution is $y(t) = \frac{1}{2}(t^2 - 1)$.

We have compared our method with the DPS continuous extension of DOPRI5(4) given by Shampine [9] whose order and computational cost is the same as our method. Thus a code which uses our continuous DOPRI5(4) and the DPS pair has been written. (Note that with both formulas the code selects the same gridpoints). Since an important application of these methods is to produce accurate output at any point of the integration interval, for each method and problem we have computed the max-norm of the global error at ten equally spaced intermediate points $t_{n,i} = t_n + \frac{1}{10} i h_n$, $i = 1, \ldots, 10$, between consecutive steps t_n and $t_{n+1} = t_n + h_n$ chosen by the integrator. Denoting by $e(t)$ the global error at the point t, we take as measure of the error for a method and a given scalar problem the quantity

$$R = \max_{n \geq 0} \left\{ \frac{\max\{ e(t_{n,i}) \mid i = 1, \ldots, 10 \}}{\max\{ e(t_n),\ e(t_{n+1}) \}} \right\}. \tag{20}$$

In the case of a nonscalar problem, the factor R in (20) is computed for each component of the problem. This means that we compute for each interval $[t_n,\ t_{n+1}]$ the ratio of the errors at ten equally spaced intermediate points divided by the greatest of the errors at the two ends of the interval and then we take the maximum over all integration intervals.

Note that instead of (20) Enright et al. [2] considered the ratio of the maximum global error at these interpolation points to the maximum global error at the grid points, i.e.,

$$R^* = \frac{\max\{ e(t_{n,i}) \mid i = 1, \ldots, 10,\ n \geq 0 \}}{\max\{ e(t_n) \mid n \geq 0 \}}. \tag{21}$$

The main reason for choosing the ratio (20) is that in problems where the global errors at the end (or more generally in some part) of the integration interval are large, the measure (21) with $R^* \cong 1$ permits errors of this size at any point of the integration interval.

The computations of the quantity (20) for the DPS and our method were carried out on a VAX-8300 computer in double precision with tolerances 10^{-i}, $i = 3, \ldots, 9$, and the numerical results are shown in Table 1. For each problem the first row corresponds to our method and the second one to the DPS method. The number in parentheses indicates the component of the vector of numerical approximations.

From this table it may be concluded that except in a few cases the errors of the interpolated values are of the same order as the error of the neighboring grid points. Furthermore in our experiments we have seen that for nonlinear equations our continuous extension is in general more reliable than the DPS. However, there are problems (e.g., D4(4), TOL $= 10^{-6}$) for which the ratio R is quite large. We have examined carefully these cases and we have found that these larger values of R are due to the contribution of the first terms of (20), typically, $n = 0$ or 1, so

Table 1

Problem	TOL $=10^{-3}$	TOL $=10^{-4}$	TOL $=10^{-5}$	TOL $=10^{-6}$	TOL $=10^{-7}$	TOL $=10^{-8}$	TOL $=10^{-9}$
A1	1.123	1.125	1.077	1.099	1.013	1.000	1.000
A1	1.039	1.047	1.019	1.002	1.000	1.000	1.000
A2	1.000	1.000	1.000	1.000	1.000	1.000	1.009
A2	1.000	1.000	1.000	1.000	1.000	1.000	1.000
A3	1.438	1.631	1.507	1.822	1.633	3.916	1.795
A3	1.267	2.823	1.940	2.059	2.044	5.303	1.977
A4	1.131	1.523	1.502	1.381	1.425	1.259	1.393
A4	1.132	1.664	1.467	1.229	1.428	1.460	2.587
(19)	1.000	1.000	2.262	1.449	1.441	1.723	1.000
(19)	1.108	1.380	6.232	6.319	11.620	3.590	1.220
D2 (1)	1.064	1.025	1.009	1.024	1.026	1.004	1.000
D2 (1)	1.064	1.025	1.010	1.032	1.035	1.009	1.004
D3 (1)	1.067	1.355	1.007	1.008	1.003	1.001	1.000
D3 (1)	1.067	1.216	1.007	1.086	1.267	1.357	1.395
D4 (1)	1.105	1.032	1.004	1.003	1.001	1.000	1.000
D4 (1)	1.105	1.031	1.004	1.173	1.478	1.680	1.780
D5 (1)	1.344	1.046	1.004	1.001	1.001	1.045	1.110
D5 (1)	1.344	1.046	1.049	1.311	1.748	2.008	2.151
D2 (2)	1.086	2.599	2.127	1.102	2.185	3.640	4.898
D2 (2)	1.446	10.350	7.765	4.607	5.191	8.545	11.440
D3 (2)	1.101	1.032	2.210	1.923	10.140	11.730	14.850
D3 (2)	1.117	2.652	7.957	4.167	21.680	24.990	27.100
D4 (2)	1.080	1.029	4.012	1.267	1.703	4.035	6.578
D4 (2)	1.083	1.668	12.210	3.327	3.807	8.255	9.822
D5 (2)	1.064	1.030	3.103	1.102	4.143	3.850	7.743
D5 (2)	1.066	2.425	8.298	2.810	4.924	13.400	12.410
D2 (3)	1.149	1.082	1.021	1.002	1.014	1.004	1.001
D2 (3)	1.139	1.082	1.020	1.003	1.013	1.004	1.001
D3 (3)	1.151	1.138	1.045	1.012	1.004	1.006	1.009
D3 (3)	1.158	1.142	1.043	1.010	1.005	1.009	1.011
D4 (3)	5.861	1.033	1.026	1.008	1.010	1.017	1.047
D4 (3)	5.067	1.033	1.021	1.008	1.014	1.022	1.034
D5 (3)	3.136	1.081	1.049	1.015	1.049	1.095	1.158
D5 (3)	2.597	1.080	1.049	1.018	1.027	1.069	1.123
D2 (4)	1.063	1.079	1.238	1.020	1.267	1.078	1.014
D2 (4)	1.063	1.077	2.421	1.107	2.500	1.189	1.049
D3 (4)	1.180	1.040	1.121	2.329	1.400	1.066	1.052
D3 (4)	1.180	1.040	1.087	4.192	1.858	1.289	1.219
D4 (4)	1.079	1.067	1.068	37.480	1.738	1.396	1.314
D4 (4)	1.079	1.067	1.036	65.810	2.465	1.747	1.581
D5 (4)	7.939	1.407	1.034	9.763	2.027	1.692	1.603
D5 (4)	7.939	1.407	1.040	14.780	2.625	2.066	1.910

that for $n \geqslant n_0$ (small) the ratios in the right-hand side of (20) are close to 1. This means that except in the first two or three steps, the errors of the interpolated values are indeed of the same order as the error of the neighboring grid points.

Acknowledgements

The authors thank the referees for their suggestions which improved the English a lot.

References

[1] J.R. Dormand and P.J. Prince, A family of embedded Runge–Kutta formulae, *J. Comput. Appl. Math.* **6** (1980).
[2] W.H. Enright, K.R. Jackson, S.P. Norsett and P.G. Thomsen, Interpolants for Runge–Kutta formulas. *ACM Trans. Math. Software* **12** (3) (1986) 193–218.
[3] E. Fehlberg, Klassische Runge–Kutta Formeln vierter und niedriger Ordnung mit Schrittweiten-Kontrolle und ihre Anwendung auf Wärmeleitungsprobleme, *Computing* **6** (1–2) (1970) 61–71.
[4] I. Gladwell, L.F. Shampine, L.S. Baca and R.W. Brankin, Practical aspects of interpolation in Runge–Kutta codes, *SIAM J. Sci. Statist. Comput.* **8** (3) (1987) 322–341.
[5] M.K. Horn, Fourth- and fifth-order, scaled Runge–Kutta algorithms for treating dense output, *SIAM J. Numer. Anal.* **20** (1983) 558–568.
[6] T.E. Hull, W.H. Enright, B.M. Fellen and A.E. Sedgwick, Comparing numerical methods for ODE's, *SIAM J. Numer. Anal.* **9** (4) (1972) 603–637.
[7] T.E. Hull, W.H. Enright and K.R. Jackson, User's guide for DVERK—A subroutine for solving nonstiff ODE's, Report 100, Dept. of Computer Science, Univ. of Toronto, Canada, 1976.
[8] L.F. Shampine, Interpolation for Runge–Kutta methods, *SIAM J. Numer. Anal.* **22** (5) (1985) 1014–1027.
[9] L.F. Shampine, Some practical Runge–Kutta formulas, *Math. Comp.* **173** (1986) 135–150.
[10] L.F. Shampine and H.A. Watts, Practical solution of ordinary differential equations by Runge–Kutta methods, Report SAND76-0585, Sandia National Laboratories, Albuquerque, NM, 1976.
[11] L.F. Shampine and H.A. Watts, DEPAC—Designed of a user oriented package of ODE solvers, Report SAND79-2374, Sandia National Laboratories, Albuquerque, NM, 1980.